Designing Second-Screening Experiences for Social Co-Selection and Critical Co-Viewing of Reality TV

Tom Feltwell, Gavin Wood, Scarlett Rowland, Kiel S. Long, Chris Elsden
Northumbria University
Newcastle upon Tyne, UK
first.last@northumbria.ac.uk

Phillip Brooker
University of Liverpool
Liverpool, UK
p.d.brooker@liverpool.ac.uk

John Vines
Northumbria University
Newcastle upon Tyne, UK
john.vines@northumbria.ac.uk

Pamela Briggs
Northumbria University
Newcastle upon Tyne, UK
p.briggs@northumbria.ac.uk

Julie Barnett
University of Bath
Bath, UK
j.c.barnett@bath.ac.uk

Shaun Lawson
Northumbria University
Newcastle upon Tyne, UK
shaun.lawson@northumbria.ac.uk

ABSTRACT

Public commentary related to reality TV can be overwhelmed by thoughtless reactions and negative sentiments, which often problematically reinforce the cultural stereotyping typically employed in such media. We describe the design and month-long evaluation of a mobile "second-screening" application, Screenr, which uses co-voting and live textual tagging to encourage more critical co-viewing in these contexts. Our findings highlight how Screenr supported interrogation of the production qualities and claims of shows, promoted critical discourse around the motivations of programmes, and engaged participants in reflecting on their own assumptions and views. We situate our results within the context of existing second-screening co-viewing work, discuss implications for such technologies to support critical engagement with socio-political media, and provide design implications for future digital technologies in this domain.

CCS CONCEPTS

• **General and reference** → **Empirical studies**; • **Human-centered computing** → *Human computer interaction (HCI)*; *Collaborative interaction.*

KEYWORDS

Second-screening, critical viewing, reality TV, co-viewing, live viewing

CHI 2019, May 4–9, 2019, Glasgow, Scotland, UK

ACM ISBN 978-1-4503-5970-2/19/05.
https://doi.org/10.1145/3290605.3300300

ACM Reference Format:
Tom Feltwell, Gavin Wood, Scarlett Rowland, Kiel S. Long, Chris Elsden, Phillip Brooker, John Vines, Pamela Briggs, Julie Barnett, and Shaun Lawson. 2019. Designing Second-Screening Experiences for Social Co-Selection and Critical Co-Viewing of Reality TV. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland, UK*. ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3290605.3300300

1 INTRODUCTION

In recent years, there has been a growing interest in the design of applications and systems that augment the viewing of television (TV). Work in this space has explored the design of peripheral devices in promoting discussion and reflection on programme content [18]; the design of mobile and tablet applications that support commentary and close readings of particular content by individual viewers [14]; and, more generally, how so-called second-screen experiences might augment and extend the viewing experience [12]. Such work often highlights the ways social media streams can provide a compelling backchannel to live TV broadcasts.

Live backchannel discussion of TV has been studied in many contexts, such as major sporting events [20], and also around politically-themed content [18]. The research has shown that whilst backchannel discussion might support some reflective and pluralistic dialogue [12], it may also lead to trolling and other forms of online abuse [11]. Brooker et al. [7] note that reality TV in particular facilitates predominantly "knee-jerk" or visceral reactions. In response, work has also explored the means to reconfigure and reimagine these conversation spaces, in order to provide more considered discussion. Feltwell et al. [14] suggested that tagging and coding media content via a second-screen might promote more critical reflection on programme content.

Motivated by these findings, we report on the design and evaluation of a mobile application, *Screenr* (Figure 1) built

to explore co-selection and critical co-viewing of reality TV. Screenr allows a group of viewers to collectively vote to decide upon a live television programme to watch together each week; the app then supports the critical co-viewing of that programme during its live broadcast. We evaluated Screenr in an in-the-wild study with 13 participants over a period of 4 weeks. During each week participants were asked to collectively decide upon a show to watch, to simultaneously view the show in their own homes, and to collaboratively comment upon, tag and discuss the show's content. The app was intentionally designed to support critical co-viewing and constructive reflection on TV content; we therefore deliberately chose reality TV content for the study, which is known to generate predominantly uncritical reactions in viewers [7].

We report on the findings of this study and explore ways in which the co-selection of programmes can be facilitated over a period of time, and how live, critical co-viewing can be supported and encouraged by apps such as Screenr. We contribute the following insights: the mechanisms required to coordinate co-selection of programmes for critical viewing, managing the attentional demand of critical second-screening, and the design decisions to encourage different types of critical reflection.

2 BACKGROUND

Second-screening refers to the common practice of interacting with a smartphone or other device (second-screen) whilst simultaneously watching a TV broadcast. Content on the second-screen is often unrelated to the TV programme being viewed, but it can also be directly related to the TV programme, either through the use of social media (e.g. using a show's hashtag), or with dedicated applications such as a companion app [19] designed to augment the experience. Such activities enrich the viewing experience, allowing not only the augmentation of the content on the primary screen, but also the connection of multiple viewing 'spheres' through social media and connected media [9]. Collaborative viewing, or *co-viewing*, is the process of watching a TV programme or video simultaneously with other people. Traditionally, co-viewing practices might simply have featured the co-located viewing of TV in a communal space, e.g. in homes or public viewing spaces. However, in recent years there has been extensive research and development around remote or distributed forms of co-viewing facilitated through second-screens and social networks. Doughty et al. [12] showed how Twitter acts to augment live viewing and create a non-co-located audience discussion. Anstead et al. [2] found the experience of co-viewing using a companion app led to an increased sociality with co-located viewers.

Critical Reflection and Critical TV Viewing

Within HCI there has been great interest in designing digital experiences to foster and encourage critical thinking, reflection and discussion around issues of contemporary societal concern. Baumer [3] explored the conceptual dimensions that underpin notions of *reflection* in HCI research, noting the assumption in his own and other work that reflection is intrinsically valuable as an experience. He explored the notion of *Inquiry*, and suggested design strategies to support this, such as the designation of separate spaces for inquiry to take place. DiSalvo [10] proposed that *adversarial design* - design that purposely challenges or even frustrates users' expectations - should be used to reconfigure 'everyday' digital technology experiences such as infographics and conversational systems, in ways that alter or invert the political standpoint supported in their use. Responding to DiSalvo's proposal, Gorkovenko et al. [18] outlined the design and deployment of living-room situated, networked printers (so called "social printers") that were used as a means to facilitate, create and broadcast discussion around political TV broadcasts. Focused on political engagement in the UK, the social printers produced printed paper discussion points that were created by other users in the network. The physicality of the printed discussion points contributed to the reconfiguration of a mundane activity - political discussion around a television broadcast - by subverting the typical format of the interactional resources around which that activity is oriented (cf. [10]). Hence, users of this system were confronted with a need to make sense of the new format of their information, and move from a passive to actively engaged state (cf. [26]).

Buschow et al. [8] noted that different types of TV programmes lend themselves to different types of (second screening) discussion. For instance, tweets around TV talent shows tend to focus critically on the personnel appearing in the programme; live events foster discussions more oriented to debating the production of the programme itself; political talk shows more readily enable discussions of the political details advanced therein. Doughty et al. [12], Brooker et al. [7], and Feltwell et al. [14] have also argued that there is particular value in building systems that engender more critically reflective content around reality TV.

Critical Reflection Focusing on Reality TV

Reality TV is a broad genre encompassing programmes that focus on competitions or contrived settings (e.g. Big Brother, X Factor), to those that claim, with no small degree of contestation, to document the goings-on of a (typically marginalised) social group [13]. The later type of reality TV is often upheld by producers as a type of documentary which should draw a reasoned critique from its audience. Yet it is typically presented as entertainment, and viewer-led discussions are not typically critical or reflective but tend to comprise 'surface'

Figure 1: Screenr usage in the home

commentary on people appearing in the show. Indeed, Scarborough and McCoy [23] suggested that viewers who report more moral (and, by implication, more 'critical') reactions to reality TV were less likely to actually watch it. Coupled with the essentially negative portrayals reality TV deals in, the tendency for reality TV to produce uncritical viewings in the majority of its audience is problematic. Given reality TV's primary purpose as entertainment, we are motivated by the previous work of Tremlett [25] who has noted how uncritical viewings of reality TV shows often accentuate the differences between the viewer and those on screen, leading to the entrenching of negative stereotypes and stigmatisation. In this way, reality TV can serve to undermine the lives of those that it claims to document, and problematically provides justification for their ongoing 'othering' and stigmatisation.

There is, therefore, a design opportunity for new modes of interaction and spaces for discussion around reality TV, where moral and critical reactions to reality TV can be promoted. The aim would be to encourage critical viewing that unpacked and explored underlying production decisions, the framing of a socio-political issue, or consideration of the viewer's own views and beliefs. In this area, Feltwell et al. [14] have explored different modes of interaction with second-screening tools that might elicit critical reflection on reality TV. Notably, they used 'social tagging' [1] of on-screen behaviours as a means to actively engage users in the critical viewing process. In these studies, incorporating a social element to 'tagging' encouraged taggers to consider the potential audiences for those tags. They noted that there is a balance to be found between attention-demanding 'lean-forward' activities such as tagging, and more passive 'lean-back' reading and reflection of other users' content.

Brooker et al. [7] analysed the online Twitter discussion of the UK reality TV series Benefits Street, which revealed the most prominent discussions were focused on the on-screen characters and were mostly 'knee-jerk' reactions aligned to the negative framing of the people depicted. A small proportion of the discussion was more critical, linking the content to wider socio-political issues. They suggested a design opportunity existed in the *"(re)design of social media platforms with a view to enhancing the potential to support more balanced, nuanced and reflective everyday socio-political talk."* [7].

3 DESIGNING FOR CRITICAL CO-VIEWING

Motivated by the previous work discussed above, we set out to design a system to support the critical co-viewing of 'everyday' reality TV programmes. Previous studies [14] have focused on encouraging critical reflection in a second-screening application without exploring the social aspect presented by co-viewing with others, facilitated through a second-screening app. Therefore, we were motivated to design a system that could encourage critical co-reflection as part of a co-viewing experience. To this end, we designed and implemented a mobile second-screening app, *Screenr*, to be used for critical co-viewing of live TV programmes. Through our design, we hoped Screenr would i) promote 'closer-readings' of TV through asking viewers to summarise on-screen talk, action, and sequences through tagging, spotting and importantly, chat; and ii) to make these readings visible to other users to foster dialogue around show content. To encourage a variety of views and differences between users, we left the tagging process open to interpretation to reflect this. The following section details the specific design decisions taken to structure co-selection of programmes, and to facilitate and encourage critical discussion during live broadcast.

The Screenr interface is shown in Figure 1. The app consists of four main features: i. co-selection of TV shows for weekly co-viewing; ii. a live interface for tagging on-screen patterns and behaviours; iii. a chat interface for open, free-form discussion around the programmes; and iv. a scrapbook for private reflection.

i. Co-selection and voting: It is imperative that a mutually agreed programme be established in order for co-viewing to be successful. Yet there are considerable socio-technical challenges when designing for co-viewing of live broadcasts. People have unique, often unpredictable schedules and may need to coordinate with other household members who also have commitments. As such, coordinating multiple users to co-view at the same time around programmes that may lie outside of their usual preferences becomes a significant scheduling problem. We selected live broadcast television for Screenr as, despite increased platform choice in recent years, industry reports (see [4]) indicate that live viewing is still the most popular way to view TV in the UK, ensuring we fit into a

normative viewing environment. TV schedules are typically released 10 days before broadcast which means coordinating multiple people for co-viewing quickly and decisively. It would have been relatively simple for the research team to dictate to users a pre-determined schedule; however we wanted participants to have some control to select shows, and encourage group discussion about the topics and programme choices available each week. In Screenr we adopted an open public voting system to coordinate co-viewing. Therefore, we made use of co-selection mechanisms, such as public voting, which have been proven to be effective tools for gauging and disseminating voting preferences of a group [27].

In the app, a digital TV guide displays details of the available programmes which users can browse through, and cast a single vote. Each programme is displayed with an image, a short summary, when and which channel it will be broadcast on, and who has voted for the programme so far (see Figure 2a). Users press a button below the programme they wish to vote for. Upon voting, their vote is reflected alongside the programmes listed in the TV guide, e.g. "User A has voted for this". Voting is open between Sunday until Wednesday morning; these days were chosen so as not to overlap with the predominant reality TV broadcast schedule (Wednesday - Friday). On Wednesday voting is automatically closed and the programme with the most votes is selected for viewing that week, with the programme selected randomly when there is a tie in votes. All users are emailed with the outcome of the vote and the selected programme's details. A log of each user's vote is publicly displayed on the home screen (Figure 2a), along with a tally of votes for each programme. Both of these mechanisms intended to expose, amongst others, the scheduling preferences of the group as a whole, and provide a degree of steering towards a mutual programme to co-view.

ii. Live interface: This is used when watching live broadcasts, and functions as a textual tagging interface for critical viewing, building on the work of [14], where tagging provides a sense-making process for the live programme. There are three key functionalities available to users: *Tagging*, *Spotting*, and *Importing* (Figure 2d and e). *Tagging* allows users to create short tags (up to 40 characters) that encapsulate an on-screen behaviour or pattern in the programme. Tags created by the user are displayed in a grid with a counter beside. *Spotting* involves tapping on a tag to mark that it has occurred in the programme and at that point in time. Spotting can be performed on any tag within the user's live interface. The amount of times a tag has been spotted is displayed by a counter alongside the tag. *Importing* allows users to copy another user's tag and bring it into their own live interface. Figure 2e shows the real-time list of all other users' tags that are available for import. Imported tags can then be used in the same way as the user's own tags, and are denoted by a purple background, compared to a green background for the user's own tags. If a user creates a tag that already exists, the existing tag is imported into the user's live interface. Tags can be deleted from the live interface at any time, and re-added should a user wish.

iii. Chat interface: In order to capture discussion in a variety of ways, we designed both a public discussion space, and private reflection space within the app. All public discussion in the app takes place in a single location, the chat interface (see Figure 2c). This is an instant-messaging style interface that allows all users to share messages with each other as part of a single, continual conversation thread.

iv. Scrapbook: While our focus is on collaborative viewing and discussion, the scrapbook allows the users to share their private thoughts with the research team, capturing any comments a user may not wish to share with the group.

Throughout the application, rather than displaying user's names, pseudonyms were used. As per prior work [14, 18], we expected this would allow users to contribute to discussion around potentially sensitive and politically charged topics without revealing their identities. At the same time, the pseudonym provided them with a unique moniker with which to be identified by other participants in subsequent conversations. Pseudonyms were selected at random from a list of animal names. A *Home* screen features as the central hub for the app (see Figure 2a) which shows recent system activity, such as user votes, vote opening, and vote closing. This is accompanied by a display of the current vote counts. The Home screen displays contextual reminders to users who have not voted, and it displays the details of the voted programme following the vote closing. Automated email reminders are also sent to notify users when new programmes are released, when they have not voted, and when voting has closed.

Screenr Weekly Cycle

Screenr operates on a two-part weekly cycle running from Sunday to Saturday, detailed below.

Part 1: Voting to watch a programme. On Sunday, new programmes become visible within the app. The voting system also becomes open simultaneously, automatically closing on Wednesday morning. Once visible, users are required to cast a single, non-transferrable vote. They do this by browsing the TV guide (Figure 2b), and hitting the vote button beside their desired option. At this point it is our intention that users will reflect on the programmes available in terms of which they think would make a good critical co-viewing experience for the group. After voting, users are prompted to discuss their voting choice in the chat interface with the rest of the group or in their own personal scrapbook. The winning programme is decided when voting closes, with

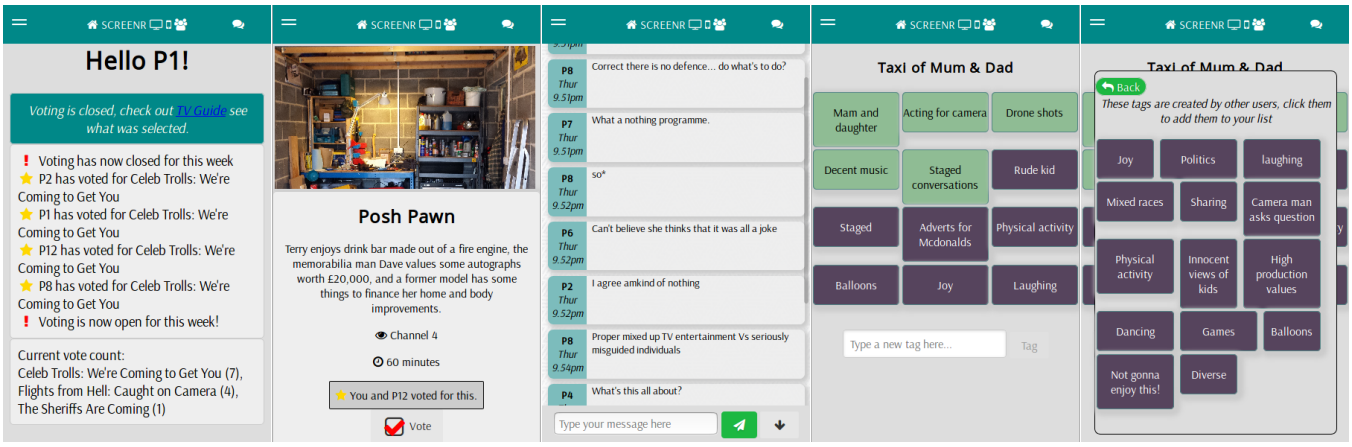

Figure 2: a) Home screen b) TV Guide for voting on programmes c) Chat interface d) Live interface e) Importing tags interface

users notified by email the outcome of the vote, reminding them what and when they will be live viewing.

Part 2: Live viewing of the selected programme. This part starts at the date and time of the selected programme. Users are directed to use the *Live Viewing* interface (Figure 2d), where they are able to create, import, and spot tags. Participants are requested to switch their TV to the correct channel. When other users create tags, a small message animates on the live interface ("New tags!") to notify the user. Users can press on this message to view the latest tags, and import any to their live interface should they chose. Throughout the programme, users are able to import, create, and spot tags in their interface. Users are able to switch to the Chat interface at any time should they wish to participate in discussion.

Technical Implementation

Screenr is implemented as a web app, primarily to avoid cross-platform compatibility issues and to make it available on a wide variety of mobile devices. The front-end of the application was created using the AngularJS framework, which incorporates HTML, CSS and JavaScript. The back-end of the system was built with Node.js, using MongoDB for database services. An API was created with Node.js and Restify, which allowed the front-end to retrieve and update information via HTTP requests.

4 STUDY DESIGN

To evaluate Screenr, and in line with the recommendations of previous work [2, 18], we conducted an in-the-wild study to explore how co-selection and critical co-viewing would work in the home environment. Notably, an in-the-wild trial would allow participants to become familiar with the app, as well as vote on and watch programmes from actual TV schedules. The study was designed to fit a normative TV viewing schedule in the UK context, using widely available channels at prime time. Participants used their own devices in their own homes with little interference from the research team.

At the commencement of the study, participants were asked to individually attend a session to orient them to the study, as well as ensure the Screenr app worked correctly on their device. During the study, participants were asked to use Screenr each week to vote for a programme and then subsequent watching of the selected programme at the allotted time. This was done each week for four weeks. In the third week of the study, participants were contacted by telephone for a 15 minute interview to ensuring they were having no technical problems with the app and to check that they were engaging with the process. In the final week of the study, three focus groups were conducted each using between 3-5 participants per group, lasting approximately one hour each. Each focus group was divided into three parts. The first part concerned the overall experiences of the participants, such as their experiences of co-viewing and which parts of the app each participant preferred to use. The second part utilised packs of physical "decks of cards" made up of all the programmes available to vote on throughout the trial, through which the researchers led participants through each week of voting and asked them to reflect on their voting choice, as well as discuss the programmes as a group. The final stage utilised printed tag clouds from each co-viewing session. Participants were asked to talk about their own co-viewing experience, and the tags they created and saw throughout the programme.

Participants

We worked with a recruitment company to recruit participants, and following the early withdrawal of one participant, the study involved 13 participants. Our inclusion criteria was

for participants to be over the age of 18, owning a recent smartphone, and regularly watching TV between 6pm and 11pm during the week. All participants lived in the north-east region of the UK to ensure they could travel into Newcastle upon Tyne for the initial interview and final focus group. Participant ages ranged from 25 to 55, with eight identifying as female and seven as male. The majority of participants were experienced in using common technology such as smartphones, laptops, and desktop computers. All members of the group owned a recent smartphone (< 3 years old), were familiar with its use, were daily viewers of TV, and regularly watched reality TV programmes. Over the duration of the four week study, the study involved participants giving up 12 hours of their time. Participants were compensated for this, with a total of £150, at an hourly rate of £12.50, based on suggested best practice [15]. To our knowledge none of the participants knew one another. At the start of the study, participants completed a 15 minute interview in order to understand their experiences of second-screening, reality TV, and othering and stigmatization on TV. They varied on a number of dimensions and these are captured below:

Reality TV viewing practices. All participants had seen reality TV. Some participants (P4, P5, P6, P7, P8, P11, P13) were enthusiastic viewers of reality TV, others (P1, P3, P9, P10, P12) were occasional viewers: *"I went off them for a little while, then I seemed to be watching more and more again"* [P10]; *"Big Brother, Love Island, I don't watch any of that, but I'll watch all the police things"* [P3]. A third group were reluctant viewers and watched simply because another family member had chosen it: *"Not necessarily through choice, it's cause what the wife's got on the TV"* [P2].

Critical Viewing. The majority of participants were relatively uncritical viewers of reality TV and would be drawn to the programmes because they provided entertainment that was easy to watch. However, a couple were critical of the genre in general: *"I don't see much constructive in it [...] poking the hornet's nest to make something happen"* [P10], or claimed a more thoughtful approach to viewing: *"I am quite analytical when I'm watching these programmes. They may be trying to portray someone as particularly violent and superficially you might firstly see the aggression in that behaviour, but I'm trying to unpick anything else they've done."* [P8].

Second-Screening. All participants, except one, had previously engaged in second-screening unrelated to the primary screen (e.g. browsing social media or conduct unrelated Internet searches). Furthermore, the majority of participants had used their second-screens in relation to the primary screen and these activities included fact-checking and using social media to view and/or join in a conversation or debate about a particular show. One participant used group messaging between friends as a back-channel for a specific programme: *"Between friends I use WhatsApp [...] That just seems to have got everyone involved."* [P1].

Choosing Programmes for Screenr

The initial criteria for selection was that programmes should be broadcast in the evening, to broadly match our participant's schedules, and that programmes should be reality TV genre. In order to reduce votes becoming too thinly spread and to provide steering for the co-viewing process, a maximum of four programmes were offered to choose between each week. The reality TV genre offered a wide range of programme topics available for critical co-viewing each week. All programmes were selected from the UK government run Freeview service [16], which includes the most commonly viewed TV channels [5].

Data Collection and Analysis

An array of usage data was collected within the Screenr system. All chat messages were recorded, including the author username and time and date stamp. Similarly, entries to the scrapbook and programme suggestions were recorded. Within the tagging interface, the following information was recorded: creation of new tags, *importing* someone else's tag to your tagging interface and *spotting* a tag (own or imported). Votes and voting choice each week were also collected for each user. In total this yielded 378 chat messages, 897 tags, 1105 imports of tags, 775 spotting instances, and 50 votes.

Each entrance interview, telephone interview, and focus group was audio recorded, and subsequently transcribed. An inductive thematic analysis method was used to analyse these transcripts, as described by [6]. Researcher 1 initially identified codes within the data, with Researcher 2 separately coding a subset of the data. Both researchers discussed their codes and, upon agreement, Researcher 1 recoded the data. Both researchers then collaboratively clustered the coded data into a set of themes.

5 FINDINGS

Our findings are divided into three sections. First, to give a sense of the system use, we give a detailed example of usage during one of the live viewing sessions. Second, we go on to describe the overall patterns of use across the duration of the trial. Finally, we discuss the outcomes of our thematic analysis, describing insights into the ways that Screenr provoked critical reflection as well as some of the social processes at play. We use an anonymised notation throughout for participants (e.g. P1), with focus group sessions being noted as FG1 through 3, and weeks abbreviated to Wk1 through 4.

Table 1: Screenr usage statistics over four weeks

	Wk1	Wk2	Wk3	Wk4
Total votes	11	13	13	12
Tags Created	297	110	299	191
Tags Imported	384	110	443	168
Tags Spotted	318	59	182	216
Chat Messages	78	42	136	122

Screenr Usage Example

We first describe an example of second-screening, taken from week three of the trial, in order to demonstrate how the system works overall. *The Taxi of Mum and Dad* was a reality TV programme aired in the UK on Channel 4 in August 2017. Billed as *"eavesdropping on conversations between parents and teens"*. The programme received 7 votes and thus was chosen to be watched by the majority of the group. The programme depicted conversations of 8 families in cars from fixed cameras pointing inwards towards the car's occupants. The programme was fast paced, showing a conversation from one family for approximately 10 minutes, with an overarching story linking the conversations throughout the programme. Two minutes before broadcast, one participant created the tag *"Sitting waiting"* [P5]. Once the programme got underway, some participants used the chat interface to reflect on their choice of programme: *"I've got this programme wrong. I thought was going to be about mam & dad taxing [sic] their kids everywhere"* [P9, Chat].

In one scene, a number of on-screen characters are seen eating fast food in the car. Within seconds of this, two users had created tags: *"Everyone likes junkfood"* and *"Advert for McDonalds"*. This second tag was then imported by two other participants into their own tagging interface. This sentiment was continued in the chat: *"Have a conversation round the dining table man... not in the car having a macci dees"* [P8, Chat]. A short montage followed, in which many of the programme's characters posed to take selfies, creating a flurry of tags within seconds of each other: *"Selfies"* [P9], and *"Selfies, so hip"* [P10]. Tags were also used to identify production features, such as the tag *"Rude son/rebellious music"* [P10] created in reference to one of the characters whose argument was accompanied by rock music.

Throughout their experience, the participants questioned the production values of the programme: *"Why on earth is this set in a car?"* [P8, Chat], which was responded with: *"No idea P8 thinking the same"* [P9, Chat]. One popular tag was created by P7: *"Acting for camera"* which was imported shortly after creation by P12 and P2, with all three users spotting this tag throughout the programme. Another participant shared their feelings at intervals during the co-viewing: *"Not real life totally staged, not what I had thought it would be!"* [P6, Chat] and *"A lot of it must be put on for the cameras"* [P6, Chat]. In this example, tags described on-screen objects and patterns, with chat providing a longer form of co-discussion. The following section provides a more in-depth analysis of this kind of participant data, taken across the whole trial period.

Overall System Usage

Four rounds of voting were conducted over the four week study, with no ties. Wk1-3 participants had four programmes to vote between, with Wk4 containing three. Votes were generally spread across all programmes, with every programme receiving at least one vote each week (see Table 2. The programmes selected tended to be on Wednesday or Thursday, starting between 7pm and 9pm. In total, 13 participants used the system for a period of 4 weeks. Of those, one engaged very little with the system throughout the study only casting votes, and whilst the remaining participants used all features of the app.

Overall, usage tended to focus on Sundays, when participants were notified of new programmes released onto the app, and on Wednesday or Thursday when the chosen programme was typically aired. Outside these times, the chat interface was used to raise technical questions (*"What is the scrapbook??"* [P2, Wk1]), or reflect on the voting process itself (*"I voted for the winner at last"* [P3, Wk3]). An overview of participant activity over the four weeks can be seen in Table 1. We can see that participant engagement remained constant, if not slightly increased, over the course of the trial, with almost all users voting. The programme viewed in week

Table 2: Programmes voted on each week.

Week	Programme	
1	Old People's Home for 4 Year Olds	5
	The Secret Life of the Holiday Resort	3
	Nightmare Tenants: Get Out of My House	2
	GPs: Behind Closed Doors	1
2	Fake Britain	7
	Nightmare Tenants, Slum Landlords	2
	Body Fixers	2
	Traffic Cops	2
3	Taxi of Mum and Dad	7
	No More Boys and Girls	3
	Nightmare Tenants, Slum Landlords	2
	Don't Deport Me, I'm British	1
4	Celeb Trolls: We're Coming to Get You	7
	Flights from Hell: Caught on Camera	4
	The Sheriffs Are Coming	1

2 was 30 minutes long (instead of 60), which thus resulted in roughly half as many interactions on Screenr. Week 3 is the most prominent in terms of engagement, with the largest amount of social interactions (e.g. chat messages, importing tags). It is evident that participants differed in the extent to which they had previously engaged in critical co-viewing and this played out in their overall use of Screenr. Figure 3 shows a representation of all 13 participants who contributed to the study, showing their contribution to each aspect of the reflective process. For example, P2 predominantly imported tags and created tags, thus their engagement with co-viewing was focused on tag curation and creation, as they imported tags heavily from the rest of the group, as well as creating tags for their own and others' use. P8 was focused very heavily on chat discussion, with a small degree of importing of other users tags, meaning they used Screenr mainly for social discussion, with tags being sourced from the rest of the group rather than creating their own. P1 in the centre has contributed evenly to all aspects of Screenr. The slight alignment to the left-hand side of the graph indicates they engaged in slightly more tagging and on-screen spotting of tags than anything else.

Through these three example participants, we can see that P2 contributed heavily to the tagging process by creating and using others' tags, with P8 contributing the most messages to the chat discussion and solely using others' tags. P1 is more balanced in their contributions, and slightly more focused on the TV screen and spotting task. Participants expressed these desires to use specific parts of the app: *"I did read the conversations through the ad breaks, then at the end of the show I'd put a comment on how I felt the show was ... The chat for me was reading what other people put."* [P11, FG3].

Results of Thematic Analysis

Our analysis of the qualitative data (tags, chat logs, notes from participants and transcripts of interviews and focus groups) captures the kinds of critical reflection that resulted from Screenr use. From our data, 6 themes were constructed which we describe below.

Critique through tagging. Through analysis of tagging behaviours, we saw that participants were often critical viewers of programme content and questioned either the "claims" made by the producers or the portrayal of people in the shows. For example, the programme in Wk1 depicted a social experiment with elders and infants. In the segment that focused on statistics and the social experiment, a large number of tags were created to interrogate these claims: *"False hope?? Revert back?"* [P2, Tags], *"How long would it last?"* [P4, Tags] and *"No negative feedback"* [P10, Tags]. Critical viewing was often associated with the depiction of controversial social issues, where participants would use chat to discuss what they considered may be provocative. In week 3, two participants had a conversation about this: *"Here we go with the class divide banter"* [P8, Chat, Wk3], shortly followed by: *"Is it [the programme] othering children or their parents, or the parent child interaction?"* [Chat, P8, Wk3]. This was followed by *"Think the othering is of parents looking clueless"* [Chat, P7, Wk3]. Participants also challenged the production choices, tagging obvious visual features: *"Close up of feet?"* [P6, Tags, Wk1], and *"Culturally diverse"* [P10, Tags, Wk3]. Critique was also directed towards portrayal of people. The tag *"Assume it's illegal immig"* [P10, Tags, Wk2] was created, and then imported twice by other participants, and within 5 minutes, one participant replied directly to the tag creator: *"I agree with whoever tagged that we automatically assume it is illegal immigrants"* [P12, Chat, Wk2].

Social affirmation and influence. There was often a strong social bonding component (which we describe in more detail in 5.3.4) whereby participants affirmed each other's reflections and discussion points: *"Agreed P8, can't ignore when they're about your kids!"* [P12, Chat, Wk4]. This often took place when participants used the chat function to discuss social issues. For example, in Wk1 discussion covered old age, and state/societal responsibility: *"Scandalous... maybe it's best if some old people aren't stimulated, it costs too much"* [P8, Chat, Wk1], and *"It does make you think though if it would have the same effect if it was a council home with kids from low income/benefits attending"* [P13, Chat, Wk1]. These appeared alongside the tags: *"Policy change afoot"* [P8, Tags, Wk1], *"Society should do more"* [P2, Tags, Wk1] and *"Do OAPs need CRBs?"* [P10, Tags, Wk1].

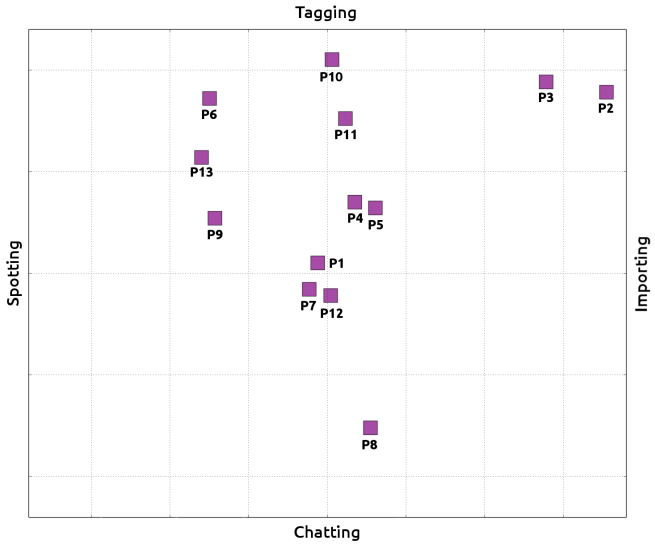

Figure 3: Balance of contribution throughout study for each of the 13 participants.

In the focus group discussions, participants recognized two sources of influence. First, aligned with comments made during the chat sessions, they acknowledged that they were sometimes manipulated by the programme producers: *"But in [the programme] the way they portrayed the northerners... the bias, the music... it categorised them"* [P7, FG2]. Others recognized that they were unable to know to what extent they were subject to manipulation: *"Is that othering on the part of the television programme, or is that just actually the reality of a lot of dodgy characters, I don't know."* [P10, FG1].

Second, the participants recognized that Screenr itself was a source of social influence, noting that they used the group as a sounding board to assess the extent to which their own views were reflected in the tagging activities of others. Some likened their use of Screenr to other social media systems such as Facebook: *"It's like a like on Facebook, using another tag"* [P7, FG2], *"I tended to look at the other tags... 'Ooo that was a good way of putting it, why didn't I put that.' My likes of other people's tags would have been in the ad breaks."* [P4, FG3]. People also recognized that the tags and chats of others could influence their own judgement: *"It was... people highlighted and I was thinking, 'Am I thinking this or am I letting people's thoughts influence me'."* [P9, FG1].

Reactionary forms of engagement. Not all tagging was as considered, however. Sometimes participants clearly displayed unthinking reactions to content, sometimes taking the form of anger directed at the people portrayed in the programme. An example is found in the chat conversation in week 3 and 4: *"Jumped up opinion of self"* [P2, Tags, Wk3], *"36 and a granny, why am I not surprised"* [P4, Chat, Wk3] and *"Vile people"* [P9, Tags, Wk4]. At other times, strong feelings of outrage would grow, partly as a function of tagging and chat as part of the second-screening process. During the fourth programme, which focused on people extensively engaged in online abuse (aka trolling), one participant discusses their own views: *"The 'extreme trolls' are so damaged... why would you do that!"* [P8, Chat, Wk4], which they follow up shortly after: *"I have totally 'othered' the trolls" ... "I'm getting so angry.." ... "I said before I wouldn't want an apology... I'd want a scalp!"* [P8, Chat, CT]. Intertwined, but not directly responding, another voices their thoughts towards their views: *"Surely these Trolls are mentally ill? 40 messages a day??"* [P7, Chat, Wk4], which is later followed by *"I'm completely othering the Trolls. Never heard of this happening, its horrendous."* [P7, Chat, Wk4]. There is a certain irony here, given that the system is developed to expose and reflect on such processes, to the fact that participants are stigmatising the people depicted in the programme and a form of social escalation takes place in the process. The implication here is that there may be behaviours such as 'flaming' or 'othering' that emerge during second-screening and we should anticipate this kind of unwelcome response.

As is typical of social media exchanges, participants could become quite heated about certain issues during the tagging and chatting process and various forms of social escalation and 'flaming' would take place. Some participants recognized this: *"it's a faceless platform, there's no worry about anybody seeing who's saying it"* [P10, FG1]; *"I think that's dangerous, and I think that's [...] why I would never fully engage with an app like Screenr [...] I would much rather have the courage and the convictions in a room full of people than the courage and convictions on a faceless app"* [P8, FG2]. As a consequence, some would carefully moderate their own input: *"I was a bit politer than I normally would be, I say what I think, but in a more guarded fashion in this group."* [P7, FG2]. Others made a purposeful decision to withhold their own views or regulate their input within the app: *"I don't want to have those sort of issues when I'm eating my dinner and relaxing [...] so I wanted to be a bit more passive, so I just ignored that one"* [P2, FG2], and *"I didn't know whether I wanted to chat with everybody on that topic [immigration], if I'm honest... I thought it could end up like... heated. I thought 'yeah let's not go there', so that's another reason why I didn't pick that one"* [P6, FG1].

Social accountability and viewing context. Participants also felt the need to be socially accountable. For example, they sometimes watched and voted on the talent programmes and in such cases would use Screenr to account for their choices, particularly if they were adrift of group norms. Sometimes they would explain the social context of the vote in terms of own their family dynamics: *"The only reason I voted for it is I've got two teenagers and I thought this is going to be real life, that's gonna be great."* [P3, FG2]. At other times they would actively involve other members of the family in the on-screen discussion: *"I was asking my husband and I was commenting on the comments that were on there and saying, 'They're all saying this', or 'Somebody's saying that, what do you think of that?' so I was involving him in the conversation. But I mean I don't think... he might've influenced what I was commenting on but only if I agreed with it and made me think about something else"* [P8, FG2].

Limitations of tag length. The word limit for tags, implemented to keep tags short and pithy, caused difficulty when trying to encapsulate on-screen behaviour or thoughts on sensitive issues: *"it was finding the tags were quite short to express really what I was trying to say a lot of the time, so a little bit more room for the tags would've probably been a bit useful."* [P10, FG1] and *"It was quite hard, trying to think of different words to sum it up, because you had a certain amount of letters that you could put in [...] how can I shorten that into a better word?"* [P4, FG3]. The character limit to tags did not present difficulty for all participants, with one method being to use tags and chat for different types of discussion:

"Tags for me, it's just summarizing what's happened in the programme [...] the [chat] was your feelings towards what was happening, what was being said" [P4, FG3] and *"With the chat you can obviously say more, so it might give more context to what the tags are aiming at. Between the two you'd probably get a good picture."* [P10, FG1]

Attention and critical co-viewing. Splitting of attention was experienced by all participants throughout the trial, with attention being directed to one of the screens based on a number of factors. One participant said that using Screenr changed their TV viewing experience from a passive viewing experience into an active second-screening experience: *"I couldn't just switch off and watch the programmes because I felt like I had to be tagging things and contributing to the discussion"* [P12, FG3]. However, the level of this active engagement was also individually moderated by the participant's interest in the programme topic: *"I felt happier when I was commenting on something I was interested in [...] watching something that I wasn't interest in, it was more of a chore."* [P8, FG2] and *"If they're heavy subjects people especially on a night time, week on week, can't be bothered to sit down and work that bit hard in watching something and trying to work out what they think about it."* [P10 FG1].

The quantity of data being generated during live viewing within Screenr could lead participants to focus on the app at the expense of the programme: *"I felt I missed quite a lot of the programme whilst I was reading what everybody else was saying and then making my own comment, and then I was like oh I've just missed like that whole 30 second segment"* [P4 FG3]. Similarly, others found they became more engaged with the discussion over time, as it changed and grew: *"At the beginning of each programme it was okay, [..] but once it started racking up, and you're trying to read what people are saying, then you're trying to think, put your bit in then somebody puts something else in it changes your"* [P3 FG2], with the attentional demand leading participants to feel unable to engage with others in the room: *"The first week my husband tried to say something and I was like 'Shh, shh!'"* [P6 FG1]. Given the attentional demand from the second-screening process, participants leveraged advert breaks, for those programmes that had them, to keep track of discussion: *"you wanted to make sure you were catching everybody else's tags. [...] somebody had said in the chat bit about not liking adverts but thank god for adverts."* [P9, FG1], *"I felt like I caught up on the conversation when the adverts were on"* [P13, FG3], and *"I used [the ad breaks] to have a rest. I just put my phone down for a couple of minutes. [...] Nobody really spoke during the ad breaks on the chat as well so, a bit of calm."* [P12 FG3].

6 DISCUSSION

We designed Screenr as a means to encourage co-selection and critical co-viewing of reality TV. Our findings reflect how we hoped participants would use Screenr, as they demonstrate users conducted close readings of the programmes viewed and engaged in a spectrum of critical discussion, whilst consuming and interacting with each others' data. Here we discuss some of the pragmatic design lessons learned from our in-the-wild co-viewing experiences before going on to consider the lessons learned about how to foster critical co-reflection.

Designing for In-The-Wild Co-Viewing

Our findings highlight a number of pragmatic considerations when designing for co-viewing in-the-wild. We unpack this in the following sections.

Provide diverse avenues for user engagement. Our results indicate a need for flexible forms of engagement as users differed in their preferences. Some predominantly used the tagging functionality - effectively "packaging" and labelling small pieces of the programme - and would use the chat sparingly and thoughtfully. Others relied on the chat interface, contributing quick comments but also sourcing tags from the group, through the importing mechanism. These show profoundly different patterns of use, yet both support co-viewing and critical reflection.

These different usage profiles also map to social roles. A user may move between a creator producing tags for the group, to a consumer importing tags, then flipping to a socialite focusing on the chat interface. This echoes existing work around social roles in tagging, such as in [24] which describes how some users may publish and produce content, whereas others may explore existing tags. We argue the need for flexibility in designing co-viewing systems to cater to a wide range of social roles and encourage diverse participation.

Open voting for co-selection. When facilitating a co-selection process, use of open and public voting mechanisms can be beneficial in a number of ways. Co-viewing of live broadcasts has numerous tensions, most notably the requirement for all users to be present at the same time, and the relatively short forecasting of TV schedules. As such, it is imperative that systems designed to facilitate the co-selection of programmes for live co-viewing leverage techniques to encourage and guide voting. At a base level, a hung vote or indecisive vote is not conducive to a self-sustaining co-selection process, and would thus require an external arbitrator, who is possibly unknown to the co-selection group. When designing for co-selection of programmes, voting as part of a group helps to build investment in the co-viewing task. As our results demonstrate, participants wanted to vote for the winning show, or reprimand their peers for the poor voting choice.

Surfacing the voting choices of the group each week also encouraged users to vote for a mutual option. Whilst this group dynamic may be useful in coordinating group co-viewing, there remains a continued pressure of individual schedules. We present the co-selection through open voting mechanism of Screenr as a means to reconcile these pressures and achieve a useful, yet participant empowering, experience.

Social facilitation. Many participants described existing second-screen behaviours wherein they would monitor social media, but often not post content themselves. When comparing those forms of engagement with the behaviours shown using Screenr, we can see that Screenr provides a relatively 'safe space' for people to air their views, supported by the small group setting. The chat allowed for a stream of commentary, allowing for quick back and forth comments which provided an element of social cohesion. Tags were less 'social' as they were anonymous, allowing users to share in a "fire and forget" way. These elements carry highly different social implications and we saw some of the disadvantages of social media borne out in Screenr, including some flaming activities, but the examples here were modest and were largely offset by expressions of group cohesion. It was interesting to note the willingness of participants to provide a social context for their viewing and we would suggest design elements that support the inclusion of other members of the family might be worth further consideration.

Fostering Critical Co-Reflection

Baumer [3] proposes that designs to encourage critical reflection should facilitate the reflective process of *Inquiry*, through creation of designated inquiry spaces, and fostering group discussion. Screenr responds to these design strategies by allowing co-viewing users to contribute to a discrete, live, social tagging interface, and a central overarching group discussion. We show that the group discussion through the chat feature fostered a wide range of critical reflection, such as on the production choices and values of programmes, the wider issues raised by the programmes, but also reflectively on participants own values and judgements.

Building group criticality. As can be seen in our results, participants are discussing othering and production techniques on their own terms, which they are not equating as negative but have still isolated this act within their conversations. This mode of discussion supported criticality among participants; this criticality was co-created together, rather than being explicitly directed by an external expert critique. As such, by reconfiguring the boundaries between 'expert' and 'TV viewer', Screenr allowed them to discuss, in their own terms, their critical reflection process. Critical viewing is often oriented towards programmes already identified as problematic (cf. [7]), however we have shown how critical tools can be turned to a variety of everyday programmes (within a specific genre), sometimes highlighting manipulations in production (Wk3), or deeper societal issues (Wk1). This wider discussion presented by the group, somewhat a 'critical collective', addresses some of the underlying issues with reality TV as a genre. Programme makers have repeatedly defended their productions as surfacing socio-political issues and highlighting these to the general public [21]. However, previous work has largely found conversations accompanying these type of programmes to be made up of negative reactions to the on-screen portrayal [7]. Whilst not unpacking the underlying issues behind the programmes, Screenr does create, through social interaction, the means to surface these wider issues. Such critical collectives may be easily formed through fan communities for reality TV (see [22]), where Screenr could be used as a means to discuss and critique the programme's content.

Engaging and disengaging with criticality. It is clear in the context of reality TV that there is a tension in relation to programme topics. In our study, when presented with voting options for programmes that centred around topical UK political issues (e.g. immigration, gender identity) some participants avoided these programmes, attributing this to not wanting to engage in any kind of discussion on these issues with others whilst watching. This raised broader questions about the applicability and frequency of use of second-screening apps for criticality. Previous work has noted how viewers may well be aware of online discussion around a programme, but purposefully chose not to engage in it [17]. We speculate that whilst 'lean-forward' engagement with second-screening can result in useful, critical engagement with the programme being viewed [26], there are some programmes where viewers would rather 'lean-back' and not engage in any discussion.

Liveness as opportunity for group criticality. Previous studies [14] have shown that users find the process of tagging during second-screening attention demanding, and within these designs a degree of passive engagement with the programme is required. The social tagging features of Screenr speak to this design consideration. As our results show, participants were able to browse tags created by others to either rephrase or reframe a topic they were interested in, or endorse and adopt a topic created by another participant by importing it to their interface. At a system level this indicates criticality can be widely sourced from the other co-viewers, which in turn reduces the attentional load for users. Furthermore, we found that participants used the advert breaks when they were available as a means to relax from the co-viewing activity. This mirrors the findings of Feltwell et al. [14] who noted that advert breaks provide a design opportunity for deeper reflection. The social aspects of Screenr - the importing of

tags, and the chat discussion - leverage this design opportunity and provide users with a space for reflection, passive or otherwise, in between and at the end of programmes.

Generalisability and Limitations

At larger scale, the quality of interaction changes, and we would imagine degrades somewhat. Specifically, an increased number of users in the chat would inevitably lead to difficulty keeping track of conversation given the increased number of messages. This study demonstrates the group size was small enough to allow for 'closer readings' of the TV content in a small, safe space. The delicate balance between creating tags, and then using chat to contextualise and explain their views would also be affected in larger groups. Likewise, the social cohesion and sense of social accountability demonstrated through the co-viewing process would reduce, leading to less meaningful voting choices and a more random viewing schedule. Programme choice was restricted in this study in order to maintain a good number of concurrent live viewers each week. At scale, with larger programme choice and more viewers, there are opportunities for machine learning to predict viewers who are likely to have similar viewing patterns, and assign them to co-viewing groups rather than restrict programme choice. Therefore we envisage that such a system would integrate as part of viewing service, for example, with many concurrent viewers (e.g. Netflix or BBC iPlayer).

The overall principles and recommendations to encourage critical co-viewing and critical reflection are applicable to topics outside reality TV, e.g. whilst co-viewing academic lectures or political speeches. One participant noted at the end of the study it would be a useful tool for co-viewing the news to collectively explore bias. It is worth noting that given the socio-political focus of reality TV, the form of critical reflection and type of reality TV watched will differ both within a country (region by region) and across countries.

7 CONCLUSION

In this work we set out to explore the co-selection and critical co-viewing of reality TV. We created a second-screening app, Screenr, to facilitate co-selection of programmes through an open voting mechanism, and critical co-viewing with tagging and chat features. Our results show Screenr's varied second-screening activities allow users to fit the system around their preferred mode of critical reflection, and as such form an ecosystem of content producers, adopters, and observers. Furthermore, we highlight a number of challenges and opportunities for the designers of critical co-viewing applications, such as the importance of providing distinct spaces for inquiry and using group discussion as a means for group sensemaking.

ACKNOWLEDGMENTS

This work was supported by the gant ES/M003574/1 "CuRAtOR: Challenging online feaR And OtheRing". CuRAtOR is funded through the Empathy and Trust in Online Communicating (EMoTICON) funding call administered by the Economic and Social Research Council in conjunction with the RCUK Connected Communities, Digital Economy and Partnership for Conflict, Crime and Security themes, and supported by the Defence Science and Technology Laboratory (Dstl) and the Centre for the Protection of National Infrastructure (CPNI). Data underpinning this paper is available from the authors.

REFERENCES

[1] Morgan Ames and Mor Naaman. 2007. Why We Tag: Motivations for Annotation in Mobile and Online Media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. ACM, New York, NY, USA, 971–980. https://doi.org/10.1145/1240624.1240772

[2] Edward Anstead, Steve Benford, and Robert J Houghton. 2014. Many-screen viewing: evaluating an olympics companion application. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video*. ACM, 103–110.

[3] Eric PS Baumer. 2015. Reflective informatics: conceptual dimensions for designing technologies of reflection. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 585–594.

[4] BBC. 2017. Telescope 2017 Report. https://www.tvlicensing.co.uk/about/telescope

[5] Broadcasters' Audience Research Board. 2017. Monthly viewing summary. http://www.barb.co.uk/viewing-data/monthly-viewing-summary/

[6] Virginia Braun, Victoria Clarke, and Gareth Terry. 2014. Thematic analysis. *Qual Res Clin Health Psychol* 24 (2014), 95–114.

[7] Phil Brooker, John Vines, Selina Sutton, Julie Barnett, Tom Feltwell, and Shaun Lawson. 2015. Debating poverty porn on Twitter: Social media as a place for everyday socio-political talk. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 3177–3186.

[8] Christopher Buschow, Beate Schneider, and Simon Ueberheide. 2014. Tweeting television: Exploring communication activities on Twitter while watching TV.

[9] Pablo Cesar, Dick CA Bulterman, and AJ Jansen. 2008. Usages of the secondary screen in an interactive television environment: Control, enrich, share, and transfer television content. In *European Conference on Interactive Television*. Springer, 168–177.

[10] Carl DiSalvo. 2012. *Adversarial design*. The MIT Press.

[11] Mark Doughty, Shaun Lawson, Conor Linehan, Duncan Rowland, and Lucy Bennett. 2014. Disinhibited abuse of othered communities by second-screening audiences. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video*. ACM, 55–62.

[12] Mark Doughty, Duncan Rowland, and Shaun Lawson. 2011. Co-viewing live TV with digital backchannel streams. In *Proceedings of the 9th European Conference on Interactive TV and Video*. ACM, 141–144.

[13] Leigh H Edwards. 2013. *The triumph of reality TV: The revolution in American television*. ABC-CLIO.

[14] Tom Feltwell, Gavin Wood, Kiel Long, Phillip Brooker, Tom Schofield, Ioannis Petridis, Julie Barnett, John Vines, and Shaun Lawson. 2017. I've been manipulated!: Designing Second Screen Experiences for

Critical Viewing of Reality TV. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2252–2263.

[15] Involve NHS National Institute for Health Research. 2012. Payment for involvement. http://www.invo.org.uk/wp-content/uploads/2012/11/INVOLVEPayment-Guiderev2012.pdf

[16] Freeview. 2018. Freeview. Video. https://www.freeview.co.uk

[17] Homero Gil de Zúñiga, Victor Garcia-Perdomo, and Shannon C McGregor. 2015. What is second screening? Exploring motivations of second screen use and its effect on online political participation. *Journal of Communication* 65, 5 (2015), 793–815.

[18] Katerina Gorkovenko, Nick Taylor, and Jon Rogers. 2017. Social Printers: a physical social network for political debates. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2269–2281.

[19] Janet Murray, Sergio Goldenberg, Kartik Agarwal, Tarun Chakravorty, Jonathan Cutrell, Abraham Doris-Down, and Harish Kothandaraman. 2012. Story-map: iPad companion for long form TV narratives. In *Proceedings of the 10th European conference on Interactive tv and video*. ACM, 223–226.

[20] Chong Oh, Sheila Sasser, and Soliman Almahmoud. 2015. Social media analytics framework: the case of Twitter and Super Bowl ads. *Journal of Information Technology Management* 26, 1 (2015), 1–18.

[21] John Plunkett. 2012. Ralph Lee: 'We don't have a malign intention, there is no agenda'. https://www.theguardian.com/media/2014/oct/26/channel-4-ralph-lee-benefits-street-immigration

[22] Cornel Sandvoss, Kelly Youngs, and Joanne Hobbs. 2015. Television Fandom in the Age of Narrowcasting: The Politics of Proximity in Regional Scripted Reality Dramas The Only Way Is Essex and Made in Chelsea. In *Popular Media Cultures*. Springer, 39–72.

[23] Roscoe C Scarborough and Charles Allan McCoy. 2016. Moral reactions to reality TV: Television viewers' endogenous and exogenous loci of morality. *Journal of Consumer Culture* 16, 1 (2016), 164–191.

[24] Jennifer Thom-Santelli, Michael J Muller, and David R Millen. 2008. Social tagging roles: publishers, evangelists, leaders. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1041–1044.

[25] Annabel Tremlett. 2014. Demotic or demonic? Race, class and gender in 'Gypsy'reality TV. *The Sociological Review* 62, 2 (2014), 316–334.

[26] Cristian Vaccari, Andrew Chadwick, and Ben O'Loughlin. 2015. Dual screening the political: Media events, social media, and citizen engagement. *Journal of Communication* 65, 6 (2015), 1041–1061.

[27] Vasilis Vlachokyriakos, Paul Dunphy, Nick Taylor, Rob Comber, and Patrick Olivier. 2014. BallotShare: An exploration of the design space for digital voting in the workplace. *Computers in Human Behavior* 41 (2014), 433–443.